CELL DELAY MODELING AND COMPARISON OF 2 ITERATIVE SCHEDULING ALGORITHMS FOR ATM INPUT - QUEUED SWITCHES

S. MOTOYAMA

Department of Telematics
School of Electrical and Computer Engineering
University of Campinas - 13081-970 Campinas, Brasil

ABSTRACT

A mathematical cell delay modeling and performance comparison of two iterative scheduling algorithms for ATM input-queued switch are carried out in this paper. The algorithms under consideration are iterative round robin with multiple classes (IRRM-MC) and iterative round robin with slip (iSLIP). By using Bernoulli arrivals a mathematical model is proposed for both algorithms. The developed model is validated by simulation. The two algorithms are compared by using Bernoulli as well as on-off arrivals. The comparison shows that input switch based on IRRM-MC algorithm is a flexible one and suitable to easily satisfy the QoS of each class of service.

1. INTRODUCTION

Several input buffered switch structures have been proposed in the literature for application in high speed ATM networks [2] - [9]. The input buffered switches have been extensively studied because of the advantages of not requiring an internal speed-up and the buffer used to queue the cells is only for an input line. These features make the input buffered switch more attractive than output buffered switch for local area backbone application where high speed characteristic is required rather than large scale feature. The main HOLB   (head of line blocking) problem that limits the throughput to 58% for FIFO type buffering [1] can be solved using many algorithms proposed in the literature to enhance the throughput, [2] - [7].

Recently, an input buffered switch with service class priority using an iterative round robin cell scheduling was proposed in [8] showing a promising switch. The scheduling algorithm used in [8] is a modified version of the algorithm presented in [9].

In this paper we propose a mathematical cell delay modeling for the input queued switch proposed in [8] and we show that the algorithm presented in [9] can be modeled as a particular case. The developed mathematical models are validated by simulation. The two algorithms are compared, and we show that by introducing a few complexity to the algorithm presented in [9], it is possible to get a flexible input queued switch suitable to satisfy easily the QoS of each service. The developed mathematical models use sources based on binomial distribution but the comparison of algorithms are also carried out using on-off type of sources.

2. ITERATIVE ROUND ROBIN ALGORITHMS

2.1 Iterative Round Robin Matching with Multiple Classes (IRRM-MC)

The switch structure proposed in [8] is presented in Fig. 1. In each input port five random access buffers (RAMs) are provided for ATM service classes (for example, CBR, rtVBR, nrtVBR, ABR and UBR). The discrimination of cells in corresponding class of services is made by switching virtual paths or virtual channels at the input of the buffers. The discrimination in service classes facilitates the achievement of the quality of service (QoS) requirement of each service. The incoming cell is stored in the buffer concerning its own service class. The cell scheduling is made in each time slot by using the control planes as shown in Fig. 1.

In each plane the cell header information is stored and is used for routing purposes. The proposed scheduling algorithm is based on iterations. The first iteration is always executed in the plane 1 and the cells from CBR service buffers are routed. The unmatched inputs in the first iteration, may be matched in the second iteration, but now this iteration is executed in the plane 2 and will route the cells from the rtVBR buffers. The plane 1 will give to the plane 2 all information about unmatched inputs and outputs in first iteration so that only these inputs and outputs are routed in second iteration. This process is performed until the fifth and last iteration in which the UBR class will be routed. The iteration process will be interrupted at any time if all inputs and outputs are matched.

Since the plane 1 is always used first in each time slot, the CBR service class will have the highest priority followed by rtVBR and UBR will have the lowest priority.

In each iteration for each priority class of service j the following three phases are used to solve the conflict among input ports (two or more cells routed to the same output port, in a time slot interval). The three phases are performed at the plane j using the input and output control units and round robin pointers ($A_{i,j}$ and $C_{i,j}$) as shown in Fig. 1.

   1. Request Phase. Each unmatched input port sends a request signal simultaneously to each output port in which it has a cell stored to route. This phase is performed at input control units (see Fig. 1) which send as a request signal only a bit.

   2. Grant Phase. (This phase is performed at output control units). If an unmatched output port i receives more than a request signal, it chooses the nearest element to the output round robin pointer $C_{i,j}$. The output port may or may not notify the input ports if their requests were accepted. The output pointer $C_{i,j}$ is incremented to the next position, to the position where the grant

signal was sent, if and only if the grant signal is accepted in third phase.

3. Accept Phase. (This phase is also performed at input control units.) If an input port *i* receives more than a grant signal, it chooses the element nearest to the input round robin pointer $A_{i,j}$. The input pointer $A_{i,j}$ that represents the highest priority element in round robin circle is incremented to the next position soon after the element is accepted .
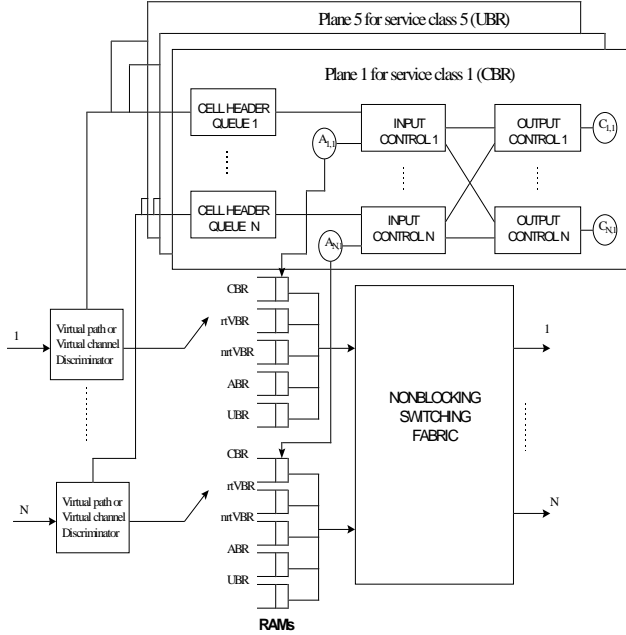

Fig. 1 - Input buffered switch with service class distinction.

2.2 Iterative Round Robin Matching with SLIP (iSLIP)

This algorithm is a modified version of PIM (Parallel Iterative Matching) presented first time in [3]. In iSLIP all three phases of the algorithm described in section 2.1 are used but without priority scheme [9]. The incoming cells are stored in only one buffer at each input and in each time slot the three phases are executed using only one plane. Thus in our analysis, iSLIP is a particular case, where the algorithm presented in section 2.1 has no priority. Also, the parallel iterative matching (PIM) algorithm presented in [3] is a particular case. In this case, the cells in phases (2) and (3) of the algorithm presented in section 2.1 are chosen randomly.

3. MODELING AND COMPARISON

3.1 Cell Delay Modeling and Performance Using Bernoulli Arrivals
3.1.1 IRRM-MC Algorithm

In the switch structure presented in [8] the cells of each class of service are distributed among N input ports, but since the cells

are chosen of each buffer in a round robin basis, for analysis the aggregate model can be considered as only one buffer for each class of service as shown in Fig. 2. The server in each time slot looks first for a cell to transmit in the CBR buffer that is the highest priority class. If there is no cell in that buffer, the server goes to the next rtVBR buffer, and so on.
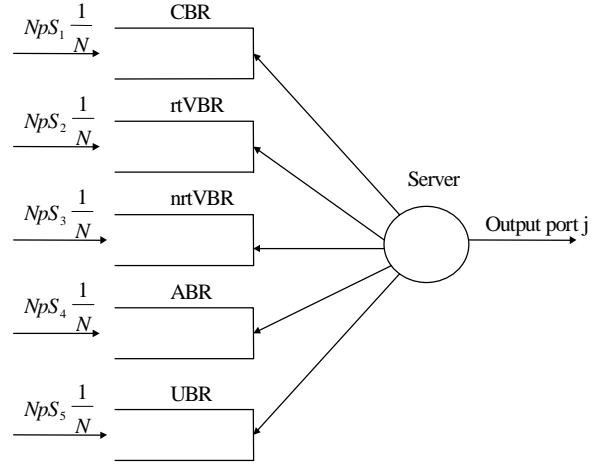

Fig. 2 - Aggregate Model for cell delay analysis.

For analysis the following assumptions are made. The number of input or output ports is N. The probability of a cell arrival in a slot is p. The traffic percentage of service class i is S_i and $\sum_i S_i = 1$. The probability of a cell at input port being routed to a particular output port is 1/N, thus the traffic at each aggregate buffer is $NpS_i \dfrac{1}{N}$. It is also assumed that there are r service classes and a generic service class of priority *h* can assume

$$h = 1, 2, 3, \ldots r$$

where r is lowest priority service class. The cells of same priority are served in a FIFO basis.

It is assumed that cell slots at inputs and outputs ports are all synchronized so that there is no residual time of service when a target cell arrives, that is, when the cells are coming at the input ports, at the same time the cells are leaving at output ports. It is also assumed that a cell is not served in the same slot in which it arrived, but in the following slots.

Using similar reasoning developed in [10] for continuous time priority queuing system with nonpreemptive service discipline, the average cell waiting time $E\{W_h\}$ for discrete time can be written as

$$E\{W_h\} = T_{slot} + \sum_{k=1}^{h} E\{T_k\} + \sum_{k=1}^{h-1} E\{T_k^{'}\} \tag{1}$$

where

T_{slot} is the time to transmit a cell.

$E\{T_k\}$ is the average time to serve all queued cells with higher priorities (h, h-1, h-2, ..., 1) when the target cell arrives.

$E\left\{T_k^{'}\right\}$ is the average time to serve all cells with higher priorities (h-1, h-2, …, 1) that arrive during the period $E\{W_h\}$ seconds and are served first before the target cell.

$E\{T_k\}$ is given by

$$E\{T_k\} = E\{m_k\}T_{slot} \qquad (2)$$

where

$E\{m_k\}$ is the average number of cells of same or higher priorities waiting for service and arrived before the target cell.

From Little's formula we obtain

$$E\{m_k\} = \frac{S_k p}{T_{slot}} E\{W_k\} \qquad (3)$$

where

$E\{W_k\}$ is the average waiting time of cells of priority k.

Thus,

$$E\{T_k\} = S_k.p.E\{W_k\} \qquad (4)$$

Similarly,

$$E\left\{T_k^{'}\right\} = S_k.p.E\{W_h\} \qquad (5)$$

Therefore,

$$E\{W_h\} = T_{slot} + \sum_{k=1}^{h} S_k.p.E\{W_k\} + E\{W_h\}\sum_{k=1}^{h-1} S_k.p \qquad (6)$$

The above equation can be solved for $E\{W_1\}$ then for $E\{W_2\}$ and by induction we get

$$E\{W_h\} = \frac{T_{slot}}{(1-p\sigma_{h-1})(1-p\sigma_h)} \qquad (7)$$

where

$$\sigma_h = \sum_{k=1}^{h} S_k \quad, \sigma_0 = 0, \quad \sigma_r = 1. \qquad (8)$$

For the case in which the cells are served in the same slots, we have

$$E\{W_h\} = \frac{T_{slot}}{(1-p\sigma_{h-1})(1-p\sigma_h)} - T_{slot}$$

$$= T_{slot}\left(\frac{1}{(1-p\sigma_{h-1})(1-p\sigma_h)} - 1\right) \qquad (9)$$

Fig. 3 shows the numerical results of the model above developed (Equation 9) compared to the results obtained by simulation of the switch using the cell scheduling algorithm presented in [8]. It was considered in the simulation that cells are served in the same slots they arrived. The percentages of load are 40%, 20%, 20%, 10% and 10%, for CBR, rtVBR, nrtVBR, ABR and UBR, respectively.

The curves of Fig. 3 show that developed model fits very well for all classes of service. For load above 90%, only for UBR class of service the model and simulation are divergent, but the load is too high to get the numerical precision in the simulation. It can be concluded that the mathematical developed model is a good one to model the cell delay performance of the input buffering switch with service class priority and using round robin cell scheduling algorithm.
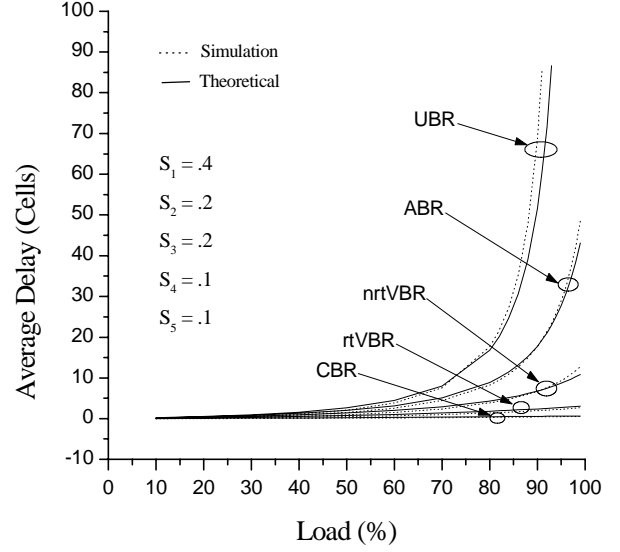

Fig. 3 - Mathematical model compared to the simulation for IRRM-MC algorithm.

3.1.2 iSLIP Algorithm

This algorithm can be considered as a particular case of the equation [9] without priority. Thus, in this case

$$\sigma_{h-1} = 0 \ , \quad \sigma_h = 1$$

$$E\{W_h\} = E\{W\} = T_{slot}\left(\frac{p}{1-p}\right) \qquad (10)$$

Fig. 4 shows the theoretical model obtained by Eq.10 compared to the simulation for iSLIP and PIM. The iSLIP and PIM algorithms have same delay performance for 4 iterations [9]. It can be noticed that the developed model fits very well to the simulation.

3.1.3 Comparison of algorithms

In the case of IRRM-MC, the cell delay for CBR class of service can be kept very low for all range of load; the iSLIP algorithm cannot guarantee this quality of service. On other hand, the UBR class may have large cell delay, but this can be avoided by controlling the percentage of load of higher priorities service classes. In IRRM-MC algorithm the cell delay can be selectively kept low according to the time constraint of each service. Thus, by introducing the priority treatment and a few implementation complexity to the iSLIP algorithm, we can obtain a high-performance ATM switch suitable to satisfy easily the QoS of each service.
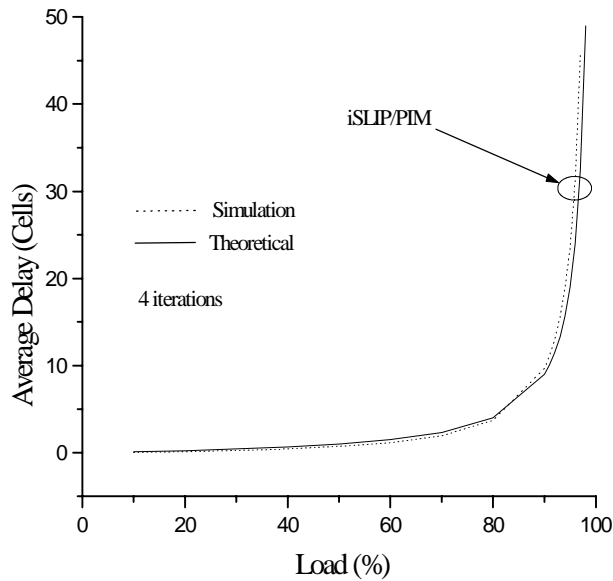
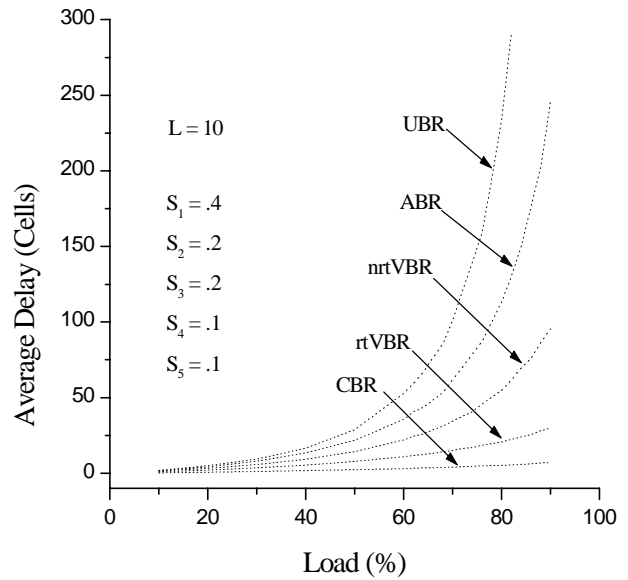**Fig. 4 -** Mathematical model compared to the simulation for iSLIP and PIM algorithms. 16x 16 switch size.

3.2 Cell Delay Performance under On-Off Traffic

In an on-off traffic, inputs alternate between active and idle periods of geometrically distributed duration. When in active period, cells arriving have the same destination port. To model this on-off source traffic, two parameters are taken into consideration: average burst length L and offered load in percentage.

In Fig. 5 obtained by simulation, it is shown a severe degradation when the switch based on IRRM-MC algorithm is submitted to bursty traffic, using L=10. For example, a load of 80% the average delays are 5.3, 20.7, 54.8 113.3 and 233.9 cells for CBR, rtVBR, nrtVBR, ABR and UBR respectively. For Bernoulli arrivals the average delays for the same load are only .34, 1.4, 3.8, 8.3 and 17.7 cells. The same degradation can be observed in iSLIP algorithm as is shown in Fig. 6.

Fig. 6 shows cell delay for iSLIP algorithm considering bursty traffic with L=10 and for 1 and 4 iterations. As it can be observed by figure the cell delay time in iSLIP algorithm for 1 iteration is very severe. For 4 iterations the algorithm behaves better although suffers severe degradation when compared to the Bernoulli arrivals. For example, for a load of 80% the average delay is 49.7 cells for on-off arrivals and only 3.7 cells for Bernoulli arrivals.

3.2.1 Comparison

Both IRRM-MC and iSLIP show severe degradation when are submitted to the on-off traffic. However, since in IRRM-MC algorithm it is possible to choose first the time constraint services, the QoS of each service can be still easily satisfied.
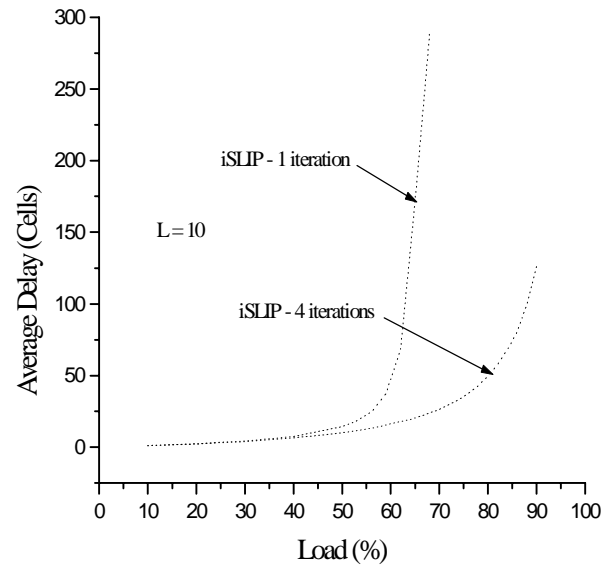

Fig. 5 - Cell delay versus load for IRRM-MC algorithm using on-off arrivals.

Fig. 6 - Cell delay versus load for iSLIP for on-off arrivals.

4. CONCLUSIONS

In this paper the cell delay modeling and performance comparison of iterative scheduling algorithms for input queued switch were carried out. By using Bernoulli arrivals it is proposed a mathematical model for both iterative round robin with multiple classes (IRRM-MC) and SLIP (iSLIP) algorithms.

The comparison of two algorithms showed that the IRRM-MC is more suitable to satisfy the QoS of each class of service. The comparison was also carried out using on-off type of arrivals. Both algorithms are severely degraded by this type of source but it can be concluded that the use of priority in classes of service is a good scheduling policy to satisfy cell delay constraint while keeping high throughput, even when the switch is submitted to a bursty traffic.

The analysis also showed that the cell delays of the two algorithms can be kept very low. However, since the IRRM-MC can serve first the time constraint classes of service, it has structure suited to satisfy easily the required quality of service (QoS).

5. REFERENCES

[1] M. Karol, M. Hluchyj, and S. P. Morgan, " Input versus output queuing in a space division switch, " *IEEE Transactions on Communications.,* vol. COM-35, pp. 1347-1356, Dec. 1987.

[2] S. Motoyama, D. W. Petr, and V. S. Frost, " Input - queued switch based on a scheduling algorithm," *Electronics Letters.*, vol. 31, no. 14, pp. 1127-1128, July 1995.

[3] T. E. Anderson, S. S. Owicki, J. B. Saxe, and C. P Tacker, "High Speed Switch Scheduling for Local Area Networks", *Proceedings of Fifth International Conference on Architectural Support for Programming Languages and Operating Systems*, pp. 98-100, Oct. 1992.

[4] H. Obara, "Optimum Architecture for Input Queuing ATM Switches", *Electronics Letters*, vol. 27, no. 7, pp. 555-557, March 1991.

[5] M. Karol, K. Eng, H. Obara, "Improving the Performance of Input-Queued ATM Packet Switches," *Proceedings of INFOCOM 92*, pp. 110 - 115, 1992.

[6] N. McKeown, P. Varaiya and J. Walrand, "Scheduling cells in an input-queued switch," *Electronics Letters*, vol. 29, no. 25, pp. 2174-2175, Dec. 1993.

[7] N. McKeown, V. Anantharam, and J. Walrand, "Achieving 100% throughput in an input-queued switch, " *IEEE Proceedings of INFOCOM'96*, San Francisco, CA, Mar. 1996.

[8] S. Motoyama, L. M. Ono, and M. C. Mavigno, "An Iterative Cell Scheduling Algorithm for ATM Input-Queued Switch with Service Class Priority" *IEEE Communications Letters*, vol. 03, no. 11, pp. 323-325, November 1999.

[9] N. McKeown, "The iSLIP Scheduling Algorithm for Input-Queued Switches" *IEEE/ACM Transactions on Networking*, vol. 07, no. 02, pp. 188-201, April 1999.

[10] J. L. Hammond, P. J. P. O´Relly, "Performance Analysis of Local Computer Networks", *Addison-Wesley Publishing Company, Chap.3,* pp.98-104, 1986.